

in accordance with a determination that the first audio stream includes the lexical trigger, generating one or more candidate text representations of the one or more utterances;

determining whether at least one candidate text representation of the one or more candidate text representations is to be disregarded by the virtual assistant;

in accordance with a determination that at least one candidate text representation is to be disregarded by the virtual assistant, generating one or more candidate intents based on candidate text representations of the one or more candidate text representations other than the to be disregarded at least one candidate text representation;

determining whether the one or more candidate intents include at least one actionable intent;

in accordance with a determination that the one or more candidate intents include at least one actionable intent, executing the at least one actionable intent;

outputting a result of the execution of the at least one actionable intent.

* * * * *